TRANSMITTAL SLIP	DATE
TO: ICS Registry	
ROOM NO.	BUILDING

REMARKS:

FROM:

ROOM NO.	BUILDING	EXTENSION

FORM NO. 241 REPLACES FORM 36-8 WHICH MAY BE USED. (47)
1 FEB 55

The Director of Central Intelligence

Washington, D.C. 20505

LOGGED

Foreign Language
Training Committee

FLTC-0015-83
3 June 1983

6 JUN 1983

Cmte 25

MEMORANDUM FOR: See Distribution

FROM: [redacted] STAT
Executive Secretary

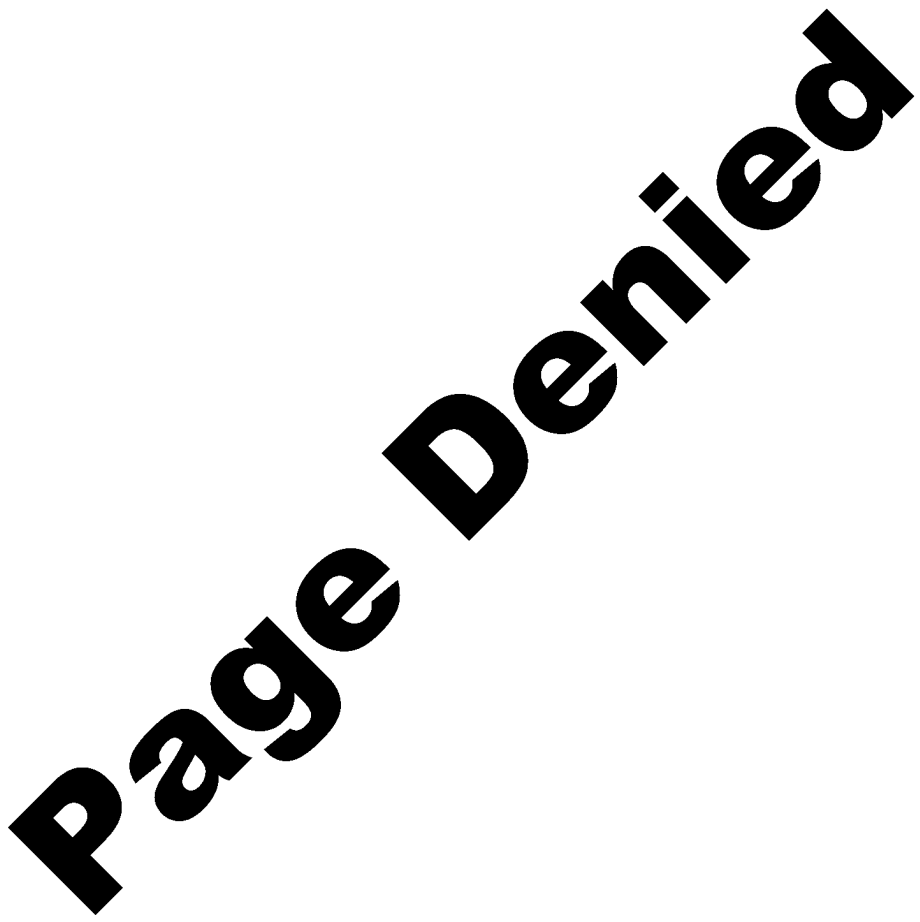
SUBJECT: The Possible Application of MAT to [redacted] STAT
[redacted] STAT

1. Attached for your information is a Memorandum for the Record covering the meeting of Intelligence Community representatives on 20 May 1983 to discuss the possible application of MAT to [redacted] STAT Also attached is a list of the names, addresses and telephone numbers of those participating in that meeting. Please advise me of any additions or corrections you wish made to the Memorandum for the Record.

2. I am scheduled to meet with [redacted] STAT of CIA/ORD, who has been designated as Chairman of a working group to explore this matter further and either she or I will keep you advised of further developments. Both [redacted] STAT and I would be grateful if you would keep us informed of any activities in this field which come to your attention and which would not ordinarily be made known to us through FLTC members, other DCI Committees, or such organizations as the Interagency Language Round Table.

[redacted] STAT

Attachments:
1. Memo for the Record
2. 1983 Jason Study for FBIS



Meeting of Intelligence Community Representatives
to Discuss the Feasibility of Utilizing Machine
Assisted Translations for Japanese S&T Documents

1. A meeting of Intelligence Community representatives concerned with and knowledgeable in the developments of MAT met under the sponsorship of the DCI's Foreign Language Training Committee (FLTC) on Friday, 20 May at the [ ] to discuss the formation of a working group to assess the feasibility of applying MAT to overtly acquire [ ] documents. A list of those attending is appended. In his opening remarks, [ ] C/FLTC, advised the participants that the Committee had been charged by the Director of the Intelligence Community Staff with coordinating the responses to the DCI Guidance (FY 1985-89) "for the CIAP and the GDIP" to "submit coordinated program initiatives designed to improve the Community's capability to produce timely and accurate translations of [ ] ... material through the ... utilization of ... machine-assisted translation technology". Moreover, the guidance also requested the CCP to accompany its 1985 budget submission with a report on the applicability of this technology to NSA's requirements. The FLTC had attempted, at this meeting, to assemble representatives of all elements of the Community interested in this problem and to constitute a working group to prepare a report to the DDCI with recommendations for possible action and funding to resolve this problem.

2. [ ] then asked [ ], Chief of FBIS' Asia Branch, to brief the group on the problem of these [ ] from the point of view of his organization which is responsible for their collection and translation. [ ] stated that FBIS had been tasked in 1980 to investigate the problem of the acquisition and [ ]

Attachment 1

[redacted]. There are at least 9000 such publications in [redacted] and FBIS STAT
has subscriptions to some 300 of them. However, it is faced with a scarcity
of competent translators of [redacted] who must be at the "5" or native fluency STAT
level to meet FBIS's rigid requirements for finished, literate translations.
An effort is now being made to supplement the translators available to FBIS in
Washington, with people recruited in [redacted] STAT
[redacted] Although the number of translators hired STAT
thus far may quadruple current output, a massive backlog of documents awaiting
screening and/or translating remains. There is, perhaps, hope in a
technological breakthrough represented by [redacted] STAT
but this technology is not yet available. Even if technological help should
be forthcoming, FBIS is faced with a most serious problem of maintaining the
morale of its human translators who fear that should FBIS move to Machine or
Machine-Assisted Translations, they would be relegated to the most boring and
menial work of editing, etc. In concluding his presentation, [redacted] STAT
emphasized that for any form of MAT suitable for his needs, machine readable
input is essential. Other participants agreed.

3. [redacted] Chairman of the Information Handling Committee STAT
(IHC), stated that the importance of [redacted] to the United STAT
States will grow "by leaps and bounds". The IHC has a charter to facilitate
the automation of all aspects of intelligence processing and will lend
whatever assistance it can to the problem of [redacted]. Even STAT
though we are successful in adapting some form of machine-assisted translation
to these documents, human assets will still be needed.

4. Mr. Bostad (FTD) reported that Dr. Thoma, president of the LATSEC
Corporation which developed the SYSTRAN system of Machine Translation utilized
by FTD, is working on a system suitable for Japanese-English pairs. Mr. Corte

(State), who is familiar with industry's needs and experimentation with MAT for [redacted], pointed out that American business has access to some 10,000 [redacted] and that one must explore the technology which could be useful in Machine assisted translation. Mr. Crawley, NISC, has been working this problem, not only for [redacted] well.

STAT STAT STAT

5. Dr. Allen Weinstein (FSI) said one must make a distinction between "translations and information capture". The latter might be well served by MAT while the former might require, in addition, the services of skilled human translators. [redacted] (ORD) agreed that machines could probably be most helpful in screening [redacted] but doubted that, at this juncture, adequate translation would be possible. [redacted] (NSA) was a bit more optomistic. Then followed a brief discussion of various systems for machine translations being evaluated or actually employed in government and industry. Mr. Corte (State) who, through his work at MIT, is familiar with industry's efforts to cope with [redacted], pointed out that the Europeans are faced with essentially the same problem as are the Americans and that it would be useful to learn what progress they have made in this field.

STAT STAT STAT STAT STAT

6. To the surprise of most, and the delight of all participants, [redacted] reported that the CIA's Directorate of Science and Technology (DDS&T) has contracted with the JASONS for a study on the applicability of MAT to [redacted] speaking for the group, welcomed this development, expressed the hope that CIA would make known to the JASONS the work presently being done in this field by the Intelligence Community, and that the JASONS would, when addressing the problem of the FBIS, take into account as well, the requirements of other elements of the Defense, Foreign

STAT STAT

Affairs, and Intelligence Communities. He advised participants that ORD has already agreed that [redacted] would be made available to chair a working group to address this problem and asked those present to advise [redacted] of their willingness to participate in such a group. [redacted] said that he would prepare a Memorandum for the Record of the meeting for review by the particpants and that he would personally contact them after meeting with [redacted] (who was away on official travel) concerning further work to be done. [redacted] thanking members for their attendance, expressed the common view that the meeting had served the most useful purpose of bringing together members of the Community who have independently been wrestling with a common problem. He trusted this would be but the beginning of a fruitful, joint cooperation.

STAT
STAT
STAT
STAT
STAT
STAT

ATTENDEES

[redacted]	C/FLTC
	ES/FLTC
	C/IHC
	ICS/PBS
	ICS/PBS/CIAP Monitor
	ICS/PBS/GDIP Monitor
	HUMINT Committee
	ICS/PPS
	CIA/ORD/Information Systems Research Division
	CIA/FBIS/C/Asia Branch
	CIA/FBIS/C/Systems Development Staff
	CIA/FBIS/Systems Development Staff
	DIA/DST/3
	NISC Translation Division, Information Management Department
	NISC Translation Division, Information Management Department
	Department of State, Foreign Service Institute, School of Language Studies
	Department of State-OES/SAT

STAT [redacted]

STAT [redacted]

STAT [redacted]

[redacted]

CIA/OTE/Language School

NSA

NSA

USAF/Foreign Technology Division (FTD)

Army ACSI

[redacted]

STAT

1983 JASON STUDY FOR FBIS

General Statement of Task

Currently, available Machine Translation (MT) systems cannot produce acceptable English translations without extensive pre-editing and/or post-editing by a human. We need to assess whether ongoing R&D activity will produce a next generation MT system which will satisfy our requirements or whether we should support some effort in a direction not now being pursued. This study will therefore consist of a survey and analysis of R&D activity (in the United States, Europe, and Japan) in artificial intelligence/natural language processing. The result of this effort will be a report and briefing identifying centers of this activity, the barriers to and shortfalls in the development of an efficient, accurate MT system, recommendations for FBIS-supported development directed toward overcoming these barriers, and the prospects of achieving such a system.

Background

FBIS currently does not have enough translators who are skilled in both the languages and the subjects of the information we collect. While we are increasing our efforts to hire and train new translators, we do not anticipate a significant increase in our translation capacity. Meanwhile, the number of foreign sources and the amount of information in foreign languages (both printed and spoken) of interest to the Intelligence Community continue to increase. Our improved collection capability must be matched by improvements in our ability to process (select, translate, and edit) the collected information.

Automation of these processing functions is a key goal of the FBIS Modernization Program. In particular, automation of the translation function would have the greatest payoff (in terms of processing through-put) if it could be achieved. At present, there are a variety of MT systems available in use which "automate," with human assistance, the translation of text from a source language to a target language. Some of these systems are simply target language word processors used by human translators. Others are claimed to be "fully automatic MT systems" because the human involvement is "pre-editing" the source text and "post-editing" the target text. The actual translation is performed by the system based on stored vocabulary and rules of grammar. All of these systems suffer from performance that is unacceptable to FBIS. That is, they produce target language products of unacceptable quality. The users of these systems must either train themselves to read and accept the MT output or they must employ a human translator in the post-edit clean-up function. Most skilled translators reject this task, claiming that is is easier to translate from the source text without the system. Consumers of FBIS products require and expect English text whose meaning and grammar are comparable to that of the source text.

Attachment 2

Plans to improve the existing MT systems by enlarging the dictionaries and refining the rules of grammar do not seem to us to be viable solutions to the problems of accurately and efficiently translating a natural (human) source language. Natural language is too ambiguous, too imprecise, too ill-defined for any dictionary/grammar based MT system to cope adequately with accurate, idiomatic translation of source language in context. A totally new approach seems required.

Assumption

We have become aware of research in Artificial Intelligence/Natural Language Processing (AI/NLP). The goal of AI/NLP (handling and understanding natural language) matches what we perceive as the significant problem in today's MT systems. We assume, therefore, that if AI/NLP techniques were applied to the translation function, the translated product would have improved quality.

Requirements

FBIS is interested solely in the translation of other languages into English. [redacted] is currently the source language of greatest interest, but an MT capability in Russian, Arabic, Chinese, and other languages of interest would be beneficial. STAT

Our goal is to eliminate the need for pre-editing but not necessarily for post-editing. Our approach is to increase the productivity (throughput rate) of the human translator by providing him with an MT system that can learn from him and require him to clean up only the unusual and "challenging" passages. A productivity increase of at least a factor of two (compared to the throughput rate of the human translator alone) does not seem unreasonable.

Our schedule permits incremental improvements in productivity. Depending on cost, we could use today an MT/human post-editor system that offered a 33% increase in productivity in translating Japanese S&T material. In two to four years, we will need productivity increases of 66% to 100%. By the end of the 80's, we will need even greater improvements for a variety of source languages.

Ancillary Issues

Although not the primary goal of this study, there are two related issues in which we may derive benefit. The first involves getting the source-language text into machine readable form. We collect foreign-language material in both printed and spoken forms. Converting this to machine readable form will be done most efficiently by foreign language optical character readers and foreign language voice detection/analysis systems.

The second related issue involves the selection function: of all the material collected, which items should be translated? If AI/NLP systems can understand the content of the source language text well enough to translate it, they should certainly be able to compare against stored criteria to perform the selection function. With a human, of course, hovering about to resolve issues of indecision.

These two issues should be kept in mind during the survey of AI/NLP research activity. Any intelligence derived, which is applicable to them, should be reported.

Questions to be Addressed

Where is research taking place in AI/NLP?

How has the field been structured?

What are the key attributes of this research?

Which attributes are critical to the development of new MT techniques?

What are the plans to develop new MT systems based on AI/NLP?

Are there existing MT systems which incorporate AI/NLP?

What are the plans to improve these systems?

What approaches should be pursued to achieve a significant improvement in the quality of current MT output?

Should FBIS support the development of an MT system employing AI/NLP?

Should FBIS acquire any existing MT system?